Decay bounds for the numerical quasiseparable preservation in matrix functions*

Stefano Massei†, Leonardo Robol‡

†Scuola Normale Superiore, Pisa,
‡Department of Computer Science, KU Leuven

Abstract

Given matrices A and B such that $B = f(A)$, where $f(z)$ is a holomorphic function, we analyze the relation between the singular values of the off-diagonal submatrices of A and B. We provide a family of bounds which depend on the interplay between the spectrum of the argument A and the singularities of the function. In particular, these bounds guarantee the numerical preservation of quasiseparable structures under mild hypotheses. We extend the Dunford-Cauchy integral formula to the case in which some poles are contained inside the contour of integration. We use this tool together with the technology of hierarchical matrices ($\mathcal{H}$-matrices) for the effective computation of matrix functions with quasiseparable arguments.

Keywords: Matrix functions, quasiseparable matrices, off-diagonal singular values, decay bounds, exponential decay, $\mathcal{H}$-matrices.

AMS subject classifications: 15A16, 65F60, 65D32, 30C30, 65E05.

1 Introduction

Matrix functions are an evergreen topic in matrix algebra due to their wide use in applications [17,20,24,26,27]. It is not hard to imagine why the interaction of structures with matrix functions is an intriguing subject. In fact, in many cases structured matrices arise and can be exploited for speeding up algorithms, reducing storage costs or allowing to execute otherwise not feasible computations. The property we are interested in is the *quasi-separability*. That is, we want to understand whether the submatrices of $f(A)$ contained in the strict upper triangular part or in the strict lower triangular part, called *off-diagonal submatrices*, have a "small" numerical rank.

Studies concerning the numerical preservation of data-sparsity patterns were carried out recently [1–3,11]. Regarding the quasiseparable structure [14,15,32,33], in [18,19,22] Gavrilyuk, Hackbusch and Khoromskij addressed the issue of approximating some matrix functions using the hierarchical format [9]. In these works the authors prove that, given a low rank quasiseparable matrix A and a holomorphic function $f(z)$, computing $f(A)$ via a quadrature formula applied to

*This work has been partially supported by an INdAM/GNCS Research Project 2016, and by the Research Council KU Leuven, project CREA/13/012, and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO.

†stefano.massei@sns.it

‡leonardo.robol@cs.kuleuven.be

the contour integral definition, yields an approximation of the result with a low quasiseparable rank. Representing A with a $\mathcal{H}$-matrix and exploiting the structure in the arithmetic operations provides an algorithm with almost linear complexity. The feasibility of this approach is equivalent to the existence of a rational function $r(z) = \frac{p(z)}{q(z)}$ which well-approximates the holomorphic function $f(z)$ on the spectrum of the argument A. More precisely, since the quasiseparable rank is invariant under inversion and sub-additive with respect to matrix addition and multiplication, if $r(z)$ is a good approximation of $f(z)$ of low degree then the matrix $r(A)$ is an accurate approximation of $f(A)$ with low quasiseparable rank. This argument explains the preservation of the quasiseparable structure, but still needs a deeper analysis which involves the specific properties of the function $f(z)$ in order to provide effective bounds to the quasiseparable rank of the matrix $f(A)$.

In this article we deal with the analysis of the quasiseparable structure of matrix functions by studying the interplay between the off-diagonal singular values of the matrices A and B such that $B = f(A)$. Our intent is to understand which parameters of the model come into play in the numerical preservation of the structure and to extend the analysis to functions with singularities.

In Section 2 we see how the integral definition of a matrix function enables us to study the structure of the off-diagonal blocks in $f(A)$. In Section 3 we develop the analysis of the singular values of structured outer products and we derive bounds for the off-diagonal singular values of matrix functions.

In Section 4 we adapt the approach to treat functions with singularities.

The key role is played by an extension of the Dunford-Cauchy formula to the case in which some singularities lie inside the contour of integration. In Section 5 we comment on computational aspects and we perform some experiments for validating the theoretical results, while in Section 6 we give some concluding remarks.

1.1 Definitions of matrix function

In [24] —which we indicate as a reference for this topic— the author focuses on three equivalent definitions of matrix function. For our purposes we recall only two of them: one based on the Jordan canonical form of the argument and the other which is a generalization of the Cauchy integral formula.

Definition 1.1. Let $A \in \mathbb{C}^{m\times m}$ and $f(z)$ be a function holomorphic in a set containing the spectrum of A. Indicating with $J = \mathrm{diag}(J_1, \ldots, J_p) = V^{-1}AV$ the Jordan canonical form of A, we define $f(A) := V \cdot f(J) \cdot V^{-1} = V \cdot \mathrm{diag}(f(J_k)) \cdot V^{-1}$ where J_k is an $m_k \times m_k$ Jordan block and

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}, \quad f(J_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \ldots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}.$$

Definition 1.2 (Dunford-Cauchy integral formula)**.** Let $f(z)$ be a holomorphic function in $\mathcal{D} \subseteq \mathbb{C}$ and $A \in \mathbb{C}^{m\times m}$ be a matrix whose spectrum is contained in $\Omega \subset \mathcal{D}$. Then we define

$$f(A) := \frac{1}{2\pi i} \int_{\partial\Omega} (zI - A)^{-1} f(z) dz. \tag{1}$$

The matrix-valued function $\mathfrak{R}(z) := (zI - A)^{-1}$ is called *resolvent*.

Suppose that the spectrum of A is contained in a disc $\Omega = B(z_0, r) := \{|z - z_0| < r\}$ where the function is holomorphic. Then, it is possible to write $f(A)$ as an integral (1) along $S^1 := \partial B(0, 1)$ for a matrix with spectral radius less than 1. In fact,

$$\frac{1}{2\pi i} \int_{\{|z-z_0|=r\}} (zI - A)^{-1} f(z) dz \quad = \quad \frac{1}{2\pi i} \int_{S^1} (wI - \tilde{A})^{-1} f(rw + z_0) dw$$

where $\tilde{A} = r^{-1}(A - z_0 I)$ has the spectrum contained in $B(0, 1)$. Given the above remark it is not restrictive to consider only the case of A having spectral radius less than 1.

Remark 1.3. In the following we will often require, besides the non singularity of $(zI - A)$, also that $(zI - D)$ is invertible along the path of integration for any trailing diagonal block D. This is not restrictive since — given a sufficiently large domain of analyticity for f — one can choose r large enough which guarantees this property. As an example, any r such that $r \geqslant \|A\|$ is a valid choice for any induced norm.

2 Off-diagonal analysis of $f(A)$

The study of the decay of the off-diagonal singular values has been investigated by [12] concerning the block Gaussian elimination on certain classes of quasiseparable matrices; in [6,7] the authors have proved fast decay properties that have been used to show the numerical quasiseparable preservation in the cyclic reduction [4,5,8,10,25].

The aim of this section is characterizing the structure of the off-diagonal blocks by means of the integral definition of $f(A)$.

2.1 Structure of an off-diagonal block

Consider the Dunford-Cauchy integral formula (1) in the case $\partial\Omega = S^1$ and A with the spectrum strictly contained in the unit disc. In this case the spectral radius of A is less than 1 and we can expand the resolvent as $(zI - A)^{-1} = \sum_{n \geqslant 0} z^{-(n+1)} A^n$.

Applying component-wise the residue theorem we find that the result of the integral in (1) coincides with the coefficient of degree -1 in the Laurent expansion of $(zI - A)^{-1} f(z)$. Thus, examining the Laurent expansion of an off-diagonal block, we can derive a formula for the corresponding block in $f(A)$. Partitioning A as follows

$$A = \begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{bmatrix} \quad \Rightarrow \quad \mathfrak{R}(z) = \begin{bmatrix} zI - \bar{A} & -\bar{B} \\ -\bar{C} & zI - \bar{D} \end{bmatrix}^{-1}$$

and supposing that the spectral radius of $\bar{D}$ is less than 1 (which is not restrictive thanks to Remark 1.3) we get

$$\mathfrak{R}(z) = \begin{bmatrix} S_{zI-\bar{D}}^{-1} & * \\ (zI - \bar{D})^{-1} \bar{C} S_{zI-\bar{D}}^{-1} & * \end{bmatrix},$$

where $S_{zI-\bar{D}} = zI - \bar{A} - \bar{B}(zI - \bar{D})^{-1}\bar{C}$ is the Schur complement of the bottom right block and $*$ denotes blocks which are not relevant for our analysis. We can write the Laurent expansion of the two inverse matrices:

$$(zI - \bar{D})^{-1} = \sum_{j \geqslant 0} z^{-(j+1)} \bar{D}^j, \qquad S_{zI-\bar{D}}^{-1} = \begin{bmatrix} I & 0 \end{bmatrix} \cdot \left(\sum_{j \geqslant 0} z^{-(j+1)} A^j \right) \cdot \begin{bmatrix} I \\ 0 \end{bmatrix},$$

where for deriving the expansion of $S^{-1}_{zI-\bar{D}}$ we used that it corresponds to the upper left block in $\mathfrak{R}(z)$.

Let $f(z) = \sum_{n\geqslant 0} a_n z^n$ be the Laurent expansion of f in S^1 and let $\mathfrak{R}(z)\cdot f(z) := \begin{bmatrix} * & * \\ G(z) & * \end{bmatrix}$, then

$$G(z) = \sum_{n\geqslant 0} a_n \sum_{j\geqslant 0} \bar{D}^j \bar{C} \cdot [I \quad 0] \cdot \sum_{s\geqslant 0} A^s z^{n-j-s-2} \cdot [I \quad 0]^t. \tag{2}$$

Exploiting this relation we can prove the following.

Lemma 2.1. *Let* $A = \begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{bmatrix}$ *be a square matrix with square diagonal blocks,* $\bar{C} = uv^t$ *and suppose that the spectrum of* A *and* $\bar{D}$ *is contained in* $B(0,1)$*. Consider* $f(z) = \sum\limits_{n\geqslant 0} a_n z^n$ *for* $|z| \leqslant 1$ *and let* $f(A) = \begin{bmatrix} * & * \\ \tilde{C} & * \end{bmatrix}$ *be partitioned according to* A*. Then*

$$\tilde{C} = \sum_{n\geqslant 1} a_n \left[u \,\middle|\, \bar{D}\cdot u \,\middle|\, \dots \,\middle|\, \bar{D}^{n-1}\cdot u\right] \cdot \left[(A^t)^{n-1}\tilde{v} \,\middle|\, \dots \,\middle|\, A^t\tilde{v} \,\middle|\, \tilde{v}\right]^t [I\ 0]^t$$

with $\tilde{v} = [I\ 0]^t v$*.*

Proof. By the Dunford-Cauchy formula, the subdiagonal block $\tilde{C}$ is equal to $\int_{S^1} G(z)dz$. By means of the residue theorem we can write the latter as the coefficient of degree -1 in (2), that is

$$\tilde{C} = \sum_{n\geqslant 1} a_n \sum_{j=0}^{n-1} \bar{D}^j uv^t \cdot [I\ 0] A^{n-j-1} [I\ 0]^t = \sum_{n\geqslant 1} a_n \sum_{j=0}^{n-1} \bar{D}^j u\tilde{v}^t A^{n-j-1} [I\ 0]^t,$$

which is in the sought form. □

Remark 2.2. The expression that we obtained for $\tilde{C}$ in the previous Lemma is a sum of outer products of vectors of the form $\bar{D}^j u$ with $(A^t)^{n-j-1}\tilde{v}$, where the spectral radii of A and $\bar{D}$ are both less than 1. This implies that the addends become negligible for a sufficiently large n. So, in order to derive bounds for the singular values, we will focus on the truncated sum

$$\sum_{n=1}^{s} a_n \left[u \,\middle|\, \bar{D}\cdot u \,\middle|\, \dots \,\middle|\, \bar{D}^{n-1}\cdot u\right] \cdot \left[(A^t)^{n-1}\tilde{v} \,\middle|\, \dots \,\middle|\, A^t\tilde{v} \,\middle|\, \tilde{v}\right]^t [I\ 0]^t \tag{3}$$

which can be rewritten as:

$$\left[u \,\middle|\, \bar{D}\cdot u \,\middle|\, \dots \,\middle|\, \bar{D}^{s-1}\cdot u\right] \cdot \left[\sum_{n=0}^{s-1} a_{n+1}(A^t)^n\tilde{v} \,\middle|\, \dots \,\middle|\, (a_s A^t + a_{s-1}I)\tilde{v} \,\middle|\, a_s\tilde{v}\right]^t [I\ 0]^t. \tag{4}$$

The columns of the left factor span the Krylov subspace $\mathcal{K}_n(\bar{D}, u) := \mathrm{Span}\{u, \bar{D}u, \dots, \bar{D}^{n-1}u\}$.

Let $p(z) := \sum_{n=0}^{s-1} a_{n+1} z^n$. Looking closely at the columns of the right factor in (4) we can see that they correspond to the so called Horner shifts (which are the intermediate results obtained while evaluating a polynomial using the Horner rule [23]) of $p(A^t)\tilde{v}$. In the following we will refer to the patterns in the factors of (4) as *Krylov* and *Horner* matrices, respectively.

3 Outer products, QR factorization and singular values

The problem of estimating the numerical rank of an outer product is addressed for example in [6], where the authors estimate the singular values of a matrix $X = \sum_{i=1}^{n} u_i v_i^*$ —where the superscript $*$ stands for the usual complex conjugate transposition— exploiting the exponential decay in the norms of the rank 1 addends. However, such an estimate is sharp only when the vectors u_i and v_i are orthogonal. In general, the singular values of X decay quickly also when the vectors u_i and/or v_i tend to become parallel as i increases. For this reason, in this work we rephrase the expression $X = \sum_{i=1}^{n} u_i v_i^*$ as $X = \sum_{i=1}^{m} \tilde{u}_i \tilde{v}_i^*$ where $\tilde{u}_i$ and $\tilde{v}_i$ are chosen as "orthogonal as possible". To this aim we study the QR decomposition of the matrices

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}, \qquad V = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

We indicate their QR decompositions as $U = Q_U R_U$ and $V = Q_V R_V$ where Q_U, Q_V are $m \times m$ and R_U, R_V have m rows and n columns.

This section is divided into five parts. In the first we study the element-wise decay in the QR factorization of Krylov matrices. In the second we show how to handle the case in which the matrix A is not diagonalizable. In the third we study the same properties for Horner matrices. In Section 3.4 we show that the singular values of a Krylov/Horner outer product inherit the decay. Finally, in Section 3.5 we derive bounds for the off-diagonal singular values of $f(A)$.

3.1 Decay in the entries of the R factor for Krylov matrices

In this section we show how to exploit the relation between Krylov subspaces and polynomial approximation [30]. More precisely, we relate the decay in the matrix R with the convergence of a minimax polynomial approximation problem in a subset of the complex plane.

The rate of convergence of the latter problem depends on the geometry of the spectrum of A. In particular, for every compact connected subset of $\mathbb{C}$ that contains the spectrum we obtain an exponent for the decay depending on its logarithmic capacity [28, 29].

In order to simplify the exposition, in this section we will assume that the matrix A is diagonalizable. However, this is not strictly required and in the next subsection we show how to relax this hypothesis.

Our approach is inspired by the one of Benzi and Boito in [1, 2], where the authors proved the numerical preservation of sparsity patterns in matrix functions. For a classic reference of the complex analysis behind the next definitions and theorems we refer to [29].

Definition 3.1 (Logarithmic capacity)**.** Let $F \subseteq \mathbb{C}$ be a nonempty, compact and connected set, and denote with G_∞ the connected component of the complement containing the point at the infinity. Since G_∞ is simply connected, in view of the Riemann Mapping Theorem we know that there exists a conformal map $\Phi(z)$ which maps G_∞ to the complement of a disc. If we impose the normalization conditions

$$\Phi(\infty) = \infty, \qquad \lim_{z \to \infty} \frac{\Phi(z)}{z} = 1$$

then this disc is uniquely determined. We say that its radius ρ is the *logarithmic capacity* of F and we write $\mathrm{lc}(F) = \rho$. Let $\Psi = \Phi^{-1}$, for every $R > \rho$ we indicate with C_R the image under Ψ of the circle $\{|z| = R\}$.

The logarithmic capacity is strictly related to the following well-known result of polynomial approximation in the complex plane.

Lemma 3.2 (Corollary 2.2 in [16]). *Let F be a Jordan region whose boundary is of finite total rotation $\mathcal{V}$ and of logarithmic capacity ρ. If $f(z)$ is an analytic function on $\mathbb{C}$ then $\forall r > \rho$ and any integer $i \geqslant 0$ there exists a polynomial $p_i(z)$ of degree at most i such that*

$$\|f(z) - p_i(z)\|_{\infty,F} \leqslant \frac{M(r)\mathcal{V}}{\pi(1-\frac{\rho}{r})} \left(\frac{\rho}{r}\right)^{i+1}.$$

with $M(r) := \max_{C_r} |f(z)|$.

In order to exploit Lemma 3.2 in our framework we need to introduce some new constants related to the geometry of the set F.

Definition 3.3. Given $F \subseteq \mathbb{C}$ compact, connected with $\mathrm{lc}(F) = \rho \in (0,1)$ we indicate with R_F the quantity

$$R_F := \sup\{R > \rho : \ C_R \text{ is strictly contained in the unit circle}\}.$$

Definition 3.4. We say that $F \subset \mathbb{C}$ is *enclosed* by $(\rho, R_F, \mathcal{V}_F)$ if $\exists F'$ Jordan region whose boundary has finite total rotation[1] $\mathcal{V}_F$, $\mathrm{lc}(F') = \rho$, $R_F = R_{F'}$ and $F \subseteq F'$.

Definition 3.5. We say that $A \in \mathbb{C}^{m\times m}$ is *enclosed* by $(\rho, R_A, \mathcal{V}_A)$ if the set of its eigenvalues is enclosed by $(\rho, R_A, \mathcal{V}_A)$.

Definition 3.6. Let J be the Jordan canonical form of $A \in \mathbb{C}^{m\times m}$. Let $\mathbb{V} := \{V \in \mathbb{C}^{m\times m} : V^{-1}AV = J\}$. We define the quantity

$$\kappa_s(A) := \inf_{V\in\mathbb{V}} \| V \|_2 \| V^{-1} \|_2 .$$

We can now proceed to study the R factor of a Krylov matrix.

Theorem 3.7. *Let $A \in \mathbb{C}^{m\times m}$ be a diagonalizable matrix enclosed by $(\rho, R_A, \mathcal{V}_A)$, $\rho \in (0,1)$ and $b \in \mathbb{C}^m$. Moreover, let U be the matrix whose columns span the n-th Krylov subspace $\mathcal{K}_n(A,b)$:*

$$U = \left[\ b \ \middle|\ Ab \ \middle|\ A^2b \quad \dots \quad A^{n-1}b \ \right].$$

Then $\forall r \in (\rho, R_A)$ the entries of the R factor in the QR decomposition of U satisfy

$$|R_{ij}| \leqslant c(r) \cdot \kappa_s(A) \cdot \left(\frac{\rho}{r}\right)^i \delta^j$$

where $\delta = \max_{z\in C_r} |z|$ and $c(r) = \frac{\mathcal{V}_A}{\delta\pi(1-\frac{\rho}{r})} \cdot \|b\|_2$.

Proof. Let $QR = U$ be the QR factorization of U and $V^{-1}AV = D$ the spectral decomposition of A. Notice that the quantity $\|R_{i+1:j,j}\|_2$ is equal to the norm of the projection of u_j on the orthogonal to the space spanned by the first i columns of U, that is $\mathcal{K}_i(A,b)^\perp$. It is well-known that the Krylov subspace $\mathcal{K}_i(A,b)$ contains all the vectors of the form $p(A)b$ where p has degree at most $i-1$. In particular, we have:

$$\begin{aligned} |R_{i+1,j}| \leqslant \|R_{i+1:j,j}\|_2 \leqslant \min_{deg(p)=i-1} \|p(A)b - u_j\|_2 &\leqslant \min_{deg(p)=i-1} \|p(D) - D^{j-1}\|_2 \|V^{-1}\|_2 \|V\|_2 \|b\|_2 \\ &\leqslant \frac{M(r)\mathcal{V}_A}{\pi(1-\frac{\rho}{r})} \left(\frac{\rho}{r}\right)^i \kappa_s(A)\|b\|_2, \end{aligned}$$

where $M(r) = \max_{C_r} |z|^{j-1} = \delta^{j-1}$. □

[1] See [16, Section 2, p. 577] for the definition of total rotation.

3.2 Non diagonalizable case

The diagonalizability hypothesis can be relaxed using different strategies. We first propose to rely on a well-known result by Crouzeix [13] based on the numerical range. Then, we discuss another approach consisting in estimating the minimax approximation error on the Jordan canonical form.

3.2.1 Numerical range

In the spirit of the results found in [1], we can give an alternative formulation that avoids the requirement of diagonalizability. The price to pay consists in having to estimate the minimax error bound on a set larger than the spectrum. To be precise, we need to consider the numerical range of the matrix A.

Definition 3.8. Let A be a matrix in $\mathbb{C}^{m\times m}$. We define its numerical range $\mathcal{W}(A)$ as the set

$$\mathcal{W}(A) = \{x^*Ax \mid x \in \mathbb{C}^m, \quad \|x\|_2 = 1\} \subseteq \mathbb{C}.$$

The numerical range is a compact convex subset of $\mathbb{C}$ which contains the eigenvalues of A. When A is normal $\mathcal{W}(A)$ is exactly the convex hull of the eigenvalues of A. Moreover, it has a strict connection with the evaluation of matrix functions [13], which is described by the following result.

Theorem 3.9 (Crouzeix)**.** *There is a universal constant $2 \leqslant \mathcal{C} \leqslant 11.08$ such that, given $A \in \mathbb{C}^{m\times m}$, and a continuous function $g(z)$ on $\mathcal{W}(A)$, analytic in its interior, the following inequality holds:*

$$\|g(A)\|_2 \leqslant \mathcal{C} \cdot \|g(z)\|_{\infty,\mathcal{W}(A)}.$$

Whenever the numerical range $\mathcal{W}(A)$ has a logarithmic capacity smaller than 1 it is possible to extend Theorem 3.7.

Corollary 3.10. *Let $A \in \mathbb{C}^{m\times m}$ be such that the field of values $\mathcal{W}(A)$ is enclosed by $(\rho, R_{\mathcal{W}(\mathcal{A})}, \mathcal{V}_{\mathcal{W}(A)})$, $\rho \in (0,1)$ and $b \in \mathbb{C}^m$. Moreover, let U be the matrix whose columns span the n-th Krylov subspace $\mathcal{K}_n(A,b)$:*

$$U = \left[\ b \ \middle|\ Ab \ \middle|\ A^2b \ \middle|\ \ldots \ \middle|\ A^{n-1}b\ \right].$$

Then $\forall r \in (\rho, R_{W(A)})$ the entries of the R factor in the QR decomposition of U satisfy

$$|R_{ij}| \leqslant c(r) \cdot \left(\frac{\rho}{r}\right)^i \delta^j$$

where $\delta = \max_{z\in C_r} |z|$ and $c(r) = \frac{\mathcal{C}\cdot\mathcal{V}_{\mathcal{W}(A)}}{\delta\pi(1-\frac{\rho}{r})} \cdot \|b\|_2$.

Proof. Follow the same steps in the proof of Theorem 3.7 employing Theorem 3.9 to bound R_{ij}. □

3.2.2 Jordan canonical form

An alternative to the above approach is to rely on the Jordan canonical form in place of the eigendecomposition. More precisely, we can always write any matrix A as $A = V^{-1}JV$ with J

being block diagonal with bidiagonal blocks (the so-called Jordan blocks). This implies that the evaluation of $f(J)$ is block diagonal with blocks $f(J_t)$ where $f(J_t)$ have the following form:

$$J_t = \begin{bmatrix} \lambda_t & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_t \end{bmatrix} \in \mathbb{C}^{m_t \times m_t}, \quad f(J_t) = \begin{bmatrix} f(\lambda_t) & f'(\lambda_t) & \dots & \frac{f^{(m_t-1)}(\lambda_t)}{(m_t-1)!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & f'(\lambda_t) \\ & & & f(\lambda_t) \end{bmatrix}.$$

We can evaluate the matrix function $f(A)$ by $f(A) = V^{-1}f(J)V$. One can estimate the norm $\|R_{i+1:j,j}\|_2$ as in the proof of Theorem 3.7:

$$|R_{i+1,j}| \leqslant \|R_{i+1:j,j}\|_2 \leqslant \min_{deg(p)=i-1} \|p(A)b - u_j\|_2 \leqslant \min_{deg(p)=i-1} \|p(J) - J^{j-1}\|_2 \cdot \kappa_s(A)_2 \cdot \|b\|_2 \quad (5)$$

where $p(J) = \mathrm{diag}(p(J_t))$, $J^j = \mathrm{diag}(J_t^j)$ and

$$p(J_t) - J_t^j = \begin{bmatrix} p(\lambda_t) - \lambda_t^j & p'(\lambda_t) - j\lambda_t^{j-1} & \dots & \frac{p^{(m_t-1)}(\lambda_t)}{(m_t-1)!} - \frac{j!}{(j-m_t)!(m_t-1)!}\lambda_h^{j-m_t} \\ & \ddots & \ddots & \vdots \\ & & \ddots & p'(\lambda_t) - j\lambda_t^{j-1} \\ & & & p(\lambda_t) - \lambda_t^j \end{bmatrix}. \quad (6)$$

We can rephrase (5) as a problem of simultaneous approximation of a function and its derivatives

Lemma 3.11. *Let $\mathcal{S}$ be a simply connected subset of the complex plane and suppose that $\exists z_0 \in \mathcal{S}$ such that each element of $\mathcal{S}$ can be connected to z_0 with a path of length less than 1. Let $p(z)$ be a degree i polynomial approximating the holomorphic function $f'(z)$ in $\mathcal{S}$, such that $|f'(z)-p(z)| \leqslant \epsilon$ in $\mathcal{S}$. Then there exists a polynomial $q(z)$ of degree $i+1$ with $q'(z) = p(z)$ such that*

$$|q(z) - f(z)| \leqslant \epsilon \qquad z \in \mathcal{S},$$

Proof. Define $q(z)$ as follows:

$$q(z) = f(z_0) + \int_\gamma p(z), \qquad \gamma \text{ any path connecting } z_0 \text{ and } z.$$

The above definition uniquely determines $q(z)$, and we know that it is a polynomial of degree $i+1$. Given $z \in \mathcal{S}$ choose γ a path connecting z_0 to z with length less than 1, we have:

$$|f(z) - q(z)| = |f(z_0) + \int_\gamma f'(z) - f(z_0) - \int_\gamma p(z)| \leqslant \int_\gamma |f'(z) - p(z)| \leqslant \epsilon.$$

□

If $m_{t'}$ is the maximum size among all the Jordan blocks we can find a minimax approximating polynomial for the $m_{t'}$ derivative of z^j. The above Lemma guarantees that, with the latter choice, the matrix (6) has the (i,j)-th entry bounded in modulus by $\frac{\epsilon}{(j-i)!}$ when $j \geqslant i$. An easy computation shows that both the 1 and ∞ norms of

$$T = \epsilon \begin{bmatrix} 1 & 1 & \frac{1}{2!} & \dots & \frac{1}{(m_{t'}-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!} \\ & & & \ddots & 1 \\ & & & & 1 \end{bmatrix}$$

are bounded by ϵe, where e is the Napier's constant. We then have $\|p(J) - J^k\|_2 \leqslant \|T\|_2 \leqslant \sqrt{\|T\|_1 \|T\|_\infty} \leqslant \epsilon e$. Using this relation one can prove the next result by following the same steps as in the proof of Theorem 3.7.

Theorem 3.12. *Let $A \in \mathbb{C}^{m\times m}$, $b \in \mathbb{C}^m$ and F be the convex hull of the spectrum of A. Suppose that $F \subseteq B(0,1)$ is enclosed by $(\rho, R_F, \mathcal{V}_F)$, $\rho \in (0,1)$ and indicate with $m_{t'}$ the size of the largest Jordan block of A. Moreover, let U be the matrix whose columns span the n-th Krylov subspace $\mathcal{K}_n(A,b)$:*

$$U = \left[\begin{array}{c|c|c|c|c} b & Ab & A^2b & \dots & A^{n-1}b \end{array}\right].$$

Then $\forall r \in (\rho, R_F)$ the entries of the R factor in the QR decomposition of U satisfy

$$|R_{ij}| \leqslant c(r) \cdot \kappa_s(A) \cdot \left(\frac{\rho}{r}\right)^{i-(m_{t'}-1)} \delta^j,$$

where $\delta = \max_{z\in C_r} |z|$ and $c(r) = \frac{e\cdot\mathcal{V}_F}{\delta\pi(1-\frac{\rho}{r})} \cdot \|b\|_2$.

3.3 Decay in the entries of the R factor for Horner matrices

Here, we show that the two-way decay in the R factor is shared by the right one in (4), which we have identified as Horner matrix.

Theorem 3.13. *Let $A \in \mathbb{C}^{m\times m}$ be a diagonalizable matrix enclosed by $(\rho, R_A, \mathcal{V}_A)$, $\rho \in (0,1)$ and $b \in \mathbb{C}^m$. Moreover let U be the matrix:*

$$U = \left[\begin{array}{c|c|c|c|c} a_s b & (a_s A + a_{s-1} I)b & \dots & \displaystyle\sum_{j=0}^{s-1} a_{j+1} A^j b \end{array}\right]$$

where the finite sequence $\{a_j\}_{j=1,\dots,s}$ verifies

$$|a_j| \leqslant \hat{\gamma} \cdot \hat{\rho}^j, \quad \hat{\gamma} > 0, \quad \hat{\rho} \in (0,1), \qquad j = 1, \dots, s.$$

Then the R factor in the QR decomposition of U is entry-wise bounded by

$$|R_{ij}| \leqslant c \cdot \kappa_s(A) \cdot \left(\frac{\rho}{R_A}\right)^i \hat{\rho}^{i+(s-j)}$$

where $c = \frac{\hat{\rho}\hat{\gamma}\mathcal{V}_A}{\pi(1-\hat{\rho})(1-\frac{\rho}{R_A})} \|b\|_2$.

Proof. Here we assume that $a_s \neq 0$. This is not restrictive because if $j < s$ is the largest j such that $a_{j'} = 0$ for any $j' > j$ the first $s-j$ columns of U are zero, and can be ignored. Observe that the j-th column of U is of the form $q(A)b$ where q is the polynomial defined by the coefficients a_j in reversed order, i.e.,

$$q(x) := \sum_{n=0}^{j-1} a_{s-j+1+n} x^n.$$

The subspace spanned by the first i columns of U contains all the vectors of the form $p(A)b$ where p is a polynomial of degree at most $i-1$. With the same argument used for proving Theorem 3.7 we can bound the entries of R in this way

$$|R_{ij}| \leqslant \min_{deg(p)=i-1} \|p(D) - \sum_{n=0}^{j-1} a_{s-j+1+n} D^n\|_2 \cdot \kappa_s(A) \cdot \|b\|_2.$$

Moreover

$$\begin{aligned}
\min_{deg(p)=i-1}\|p(D)-\sum_{n=0}^{j-1}a_{s-j+1+n}D^n\|_2 &= \min_{deg(p)=i-1}\|p(D)-\sum_{n=i}^{j-1}a_{s-j+1+n}D^n\|_2\\
&\leqslant \sum_{n=i}^{j-1}|a_{s-j+1+n}|\min_{deg(p)=i-1}\|p(D)-D^n\|_2\\
&\leqslant \sum_{n=i}^{j-1}\hat{\gamma}\hat{\rho}^{s-j+1+n}\min_{deg(p)=i-1}\|p(D)-D^n\|_2\\
&\underbrace{\leqslant}_{\text{Lemma 3.2}}\sum_{n=i}^{j-1}\hat{\gamma}\hat{\rho}^{s-j+1+n}\frac{\mathcal{V}_A}{\pi(1-\frac{\rho}{R_A})}\left(\frac{\rho}{R_A}\right)^i\\
&\leqslant \frac{\hat{\rho}\hat{\gamma}\mathcal{V}_A}{\pi(1-\hat{\rho})(1-\frac{\rho}{R_A})}\hat{\rho}^{s-j+i}\left(\frac{\rho}{R_A}\right)^i,
\end{aligned}$$

where we used Lemma 3.2 with $r = R_A$. □

Remark 3.14. In view of the above arguments we can rephrase Theorem 3.7 for non diagonalizable matrices. We obtain similar statements involving $\mathrm{lc}(\mathcal{W}(A))$ in place of $\mathrm{lc}(A)$ or with a shifted column decay. The same technique can be used to generalize the results of the next sections. The proofs and statements are analogous to the diagonalizable case. Therefore, we do not report them.

3.4 Decay in the singular values of Krylov/Horner outer products

3.4.1 Some preliminaries

In what follows, we indicate with Π_m the counter identity of order m:

$$\Pi_m := \begin{bmatrix} & & 1\\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{m\times m},$$

which is the matrix which flips the columns.

Due to technical reasons, we also need to introduce the following quantity.

Definition 3.15. Given $A \in \mathbb{C}^{m\times m}$ enclosed by $(\rho, R_A, \mathcal{V}_A)$ and a parameter $R \in \mathbb{R}^+$ we define

$$\Lambda(\rho, R_A, \mathcal{V}_A, R) := \frac{\mathcal{V}_A^2}{\pi^2(R-1)(1-\frac{\rho}{R_A})\sqrt{1-(\frac{\rho}{RR_A})^2}}\cdot\min_{\rho<r<R_A}\frac{1}{\delta(r)(1-\delta(r)^2)(\frac{r}{\rho}-1)\sqrt{(1-\frac{\rho^2}{r^2})}},$$

where $\delta(r) := \max\{\frac{1}{R}, \max_{C_r}|z|\}$.

3.4.2 The estimates

Now, we have all the ingredients for studying the singular values of Krylov/Horner outer products. For simplicity we state a result in the diagonalizable case, but we highlight that it is easy to recover analogous estimates for the general framework employing the techniques of Section 3.2.

Theorem 3.16. *Let $b_1 \in \mathbb{C}^m$, $b_2 \in \mathbb{C}^n$ and $A_1 \in \mathbb{C}^{m\times m}, A_2 \in \mathbb{C}^{n\times n}$ be two diagonalizable matrices enclosed by $(\rho, R_A, \mathcal{V}_A)$ with $\rho \in (0,1)$. Then for any finite sequence $\{a_j\}_{j=1,\dots,s}$ which verifies*

$$|a_j| \leqslant \hat{\gamma} \cdot R^{-j}, \quad R > 1, \quad j \in \{1, \dots, s\},$$

the singular values of

$$X = \left[\begin{array}{c|c|c|c} b_1 & A_1 b_1 & \dots & A_1^{s-1} b_1 \end{array}\right] \cdot \left[\begin{array}{c|c|c|c} \displaystyle\sum_{j=0}^{s-1} a_{j+1} A_2^j b_2 & \dots & (a_s A_2 + a_{s-1} I) b_2 & a_s b_2 \end{array}\right]^t \tag{7}$$

can be bounded by

$$\sigma_l(X) \leqslant \gamma \cdot e^{-(\alpha+\alpha')(l+1)}, \qquad \alpha = \log\left(\frac{R_A}{\rho}\right), \qquad \alpha' = \log(R),$$

where $\gamma := \hat{\gamma} \cdot \kappa_s(A_1)\kappa_s(A_2)\|b_1\|_2\|b_2\|_2 \cdot \Lambda(\rho, R_A, \mathcal{V}_A, R)$.

Proof. Consider the matrices U and V defined as follows:

$$U = \left[\begin{array}{c|c|c|c} b_1 & A_1 b_1 & \dots & A_1^{s-1} b_1 \end{array}\right], \quad V = \left[\begin{array}{c|c|c|c} a_s b_2 & (a_s A_2 + a_{s-1} I) b_2 & \dots & \displaystyle\sum_{j=0}^{s-1} a_{j+1} A_2^j b_2 \end{array}\right],$$

so that we have $X = U\Pi_s V^t$ as in Equation (7). Moreover, let (Q_U, R_U) and (Q_V, R_V) be the QR factorizations of U and V respectively. Applying Theorem 3.7 and Theorem 3.13 we get that $\forall r \in (\rho, R_A)$

$$|R_{U,ij}| \leqslant c_1(r) \cdot e^{-\eta i - \beta j} \qquad \text{and} \qquad |R_{V,ij}| \leqslant c_2 \cdot e^{-(\alpha+\alpha')i - \beta(s-j)},$$

with $\eta = \log\left(\frac{r}{\rho}\right)$, $\beta = |\log(\delta)|$, $c_1(r) = \frac{\mathcal{V}_{A_1}}{\delta\pi(1-\frac{\rho}{r})} \cdot \kappa_s(A_1) \cdot \|b_1\|_2$ and $c_2 = \frac{\hat{\rho}\hat{\gamma}\mathcal{V}_{A_2}}{\pi(1-\hat{\rho})(1-\frac{\rho}{R_A})}\kappa_s(A_2)\|b_2\|_2$.

In order to bound the singular values of X we look at those of $S = R_U \Pi_s R_V^*$. The entry (i,j) of S is obtained as the sum:

$$S_{ij} = \sum_{h=1}^{s} R_{U,ih} \cdot R_{V,j(s-h)}, \qquad |R_{U,ih} \cdot R_{V,j(s-h)}| \leqslant c \cdot e^{-\eta i - (\alpha+\alpha')j - 2\beta h},$$

where $c = c_1(r) \cdot c_2$. Summing all the bounds on the addends we obtain

$$|S_{ij}| \leqslant \frac{c}{1 - e^{-2\beta}} e^{-\eta i - (\alpha+\alpha')j}.$$

We can estimate the l-th singular value by setting the first $l-1$ columns of S to zero. Let S_l be the matrix composed by the last $m - l + 1$ columns of S. Since this matrix can be seen as the residue of a particular choice for a rank $l-1$ approximation of S we have $\sigma_l(S) \leqslant \|S_l\|_2$. The entries of S_l satisfy the relation $(S_l)_{ij} \leqslant \tilde{\gamma} e^{-(\alpha+\alpha')l} e^{-\eta i - (\alpha+\alpha')(j-1))}$ where $\tilde{\gamma} = \frac{c}{1-e^{-2\beta}}$, so we obtain:

$$\left\| \frac{e^{(\alpha+\alpha')l}}{\tilde{\gamma}} S_l \right\|_F^2 = \sum_{i=1}^{m-l} \sum_{j=1}^{n} |\frac{e^{(\alpha+\alpha')l}}{\tilde{\gamma}} (S_k)_{i,j}|^2 \leqslant \frac{e^{-2\eta}}{(1 - e^{-2\eta})(1 - e^{(-\alpha+\alpha')})}.$$

Since $\|S_l\|_2 \leqslant \|S_l\|_F$ we have $\sigma_l(S) \leqslant \frac{\tilde{\gamma} e^{-\eta}}{\sqrt{(1-e^{-2\eta})(1-e^{-2(\alpha+\alpha')})}} e^{-(\alpha+\alpha')l} = \gamma e^{-(\alpha+\alpha')l}$. □

Our final aim is to estimate the singular values of (3) by estimating the singular values of one of its finite truncations (4). In order to justify that, we need to show that the addends in (3) become negligible. Observe that the latter are outer products of two Krylov matrices in which the second factor appears in a reverse order. This means that the row-decay in its R factor has an opposite direction. In the next result we see how this fact implies the negligibility.

Theorem 3.17. *Let $U = Q_U R_U$ and $V = Q_V R_V$ be QR factorizations of $U \in \mathbb{C}^{m\times n}$ and $V \in \mathbb{C}^{m\times n}$. Let α, β and c be positive constants such that $|R_{U,ij}|, |R_{V,ij}| \leqslant ce^{-\alpha i-\beta j}$ for any i, j. Then the matrix $X = U\Pi_n V^*$ has singular values bounded by*

$$\sigma_l(X) \leqslant \gamma e^{-\alpha(l+1)}, \qquad \gamma := \frac{c^2 n e^{-(n+1)\beta}}{(1-e^{-2\alpha})}.$$

Proof. We can write $X = U\Pi_n V^* = Q_U R_U \Pi_n R_V^* Q_V^*$, so its singular values coincide with the ones of $S = R_U \Pi_{n,m} R_V^*$. The element in position (i, j) of S is obtained as the a sum

$$S_{ij} = \sum_{l=1}^{n} R_{U,il} \cdot R_{V,j(n-l)}, \qquad |R_{U,il} \cdot R_{V,j(n+1-l)}| \leqslant c^2 e^{-\alpha(i+j)-\beta(n+1)}$$

according to our hypotheses. Since the bound on the elements in the above summation is independent of n we can write $|S_{ij}| \leqslant c^2 n e^{-\beta(n+1)} e^{-\alpha(i+j)}$. The thesis can then be obtained by following the same procedure as in Theorem 3.16. □

Remark 3.18. Observe that the larger n the closer the quantity $ne^{-\beta n}$ is to 0. Therefore for sufficiently big n the resulting matrix is negligible.

3.5 Decay in the off-diagonal singular values of $f(A)$

We start with a few technical results that will make some proofs smoother.

Lemma 3.19. *Let $A^+ = \sum_{j=0}^{+\infty} A_j$ with $A_j \in \mathbb{R}^{m\times n}$ matrices of rank k and suppose that $\|A_j\|_2 \leqslant \gamma e^{-\alpha|j|}$. Then*

$$\sigma_l(A^+) \leqslant \frac{\gamma}{1-e^{-\alpha}} \cdot e^{-\alpha\frac{l-k}{k}}.$$

Proof. Note that $\sum_{j<\lceil\frac{l-k}{k}\rceil} A_j$ is at most a rank-$(l-1)$ approximation of A. This implies that

$$\sigma_l(A) \leqslant \left\| A - \sum_{j<\lceil\frac{l-k}{k}\rceil} A_j \right\|_2 = \left\| \sum_{j\geqslant\lceil\frac{l-k}{k}\rceil} A_j \right\|_2 \leqslant \sum_{j\geqslant\lceil\frac{l-k}{k}\rceil} \gamma e^{-\alpha j} =$$
$$= \gamma e^{-\alpha\lceil\frac{l-k}{k}\rceil} \sum_{j\geqslant 0} e^{-\alpha j} = \frac{\gamma}{1-e^{-\alpha}} \cdot e^{-\alpha\lceil\frac{l-k}{k}\rceil}.$$

□

Lemma 3.20. *Let $A = \sum_{i=1}^{k} A_i \in \mathbb{C}^{n\times n}$ where $\sigma_j(A_i) \leqslant \gamma e^{-\alpha j}$, for $j = 1, \ldots, n$. Then $\sigma_j(A) \leqslant \tilde{\gamma} e^{-\alpha\frac{j-k}{k}}$, $\tilde{\gamma} = \frac{k\gamma}{1-e^{-\alpha}}$.*

Proof. Relying on the SVD, we write $A_i = \sum_{j=1}^{\infty} \sigma_j(A_i) u_{i,j} v_{i,j}^*$ where $u_{i,j}$ and $v_{i,j}$ are the singular vectors of A_i and where, for convenience, we have expanded the sum to an infinite number of terms by setting $\sigma_j(A_i) = 0$ for $j > n$. This allows us to write

$$A = \sum_{i=1}^{k} A_i = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{k} \sigma_j(A_i) u_{i,j} v_{i,j}^* \right) = \sum_{j=1}^{\infty} \tilde{A}_j.$$

Observe that $\tilde{A}_j$ have rank k and $\|A_j\| \leqslant k\gamma e^{-\alpha j}$. Applying Lemma 3.19 completes the proof. □

Lemma 3.21. *Let $A, B \in \mathbb{C}^{m \times m}$ and suppose that B has rank k. Then*

$$\sigma_{j+k}(A + B) \leqslant \sigma_j(A).$$

Proof. For the Eckart-Young-Mirsky theorem $\forall j = 1, \ldots, m \; \exists \widetilde{A}$ of rank j such that $\|A - \widetilde{A}\|_2 = \sigma_j(A)$. Therefore, since $\widetilde{A} + B$ has rank less than or equal to $j + k$ we have

$$\sigma_{j+k}(A + B) \leqslant \|(A + B) - (\widetilde{A} + B)\|_2 = \sigma_j(A).$$

□

We are ready to study singular values of the matrix resulting from applying a function to a matrix. We prefer to begin by stating a simpler result which holds for matrices with spectrum contained in $B(0, 1)$ and function holomorphic on a larger disk. In the following corollaries it is shown how to adapt this result to more general settings.

Theorem 3.22. *Let $A \in \mathbb{C}^{m \times m}$ be quasiseparable of rank k and such that A and all its trailing submatrices are enclosed in $(\rho, R_A, \mathcal{V}_A)$ and diagonalizable. Consider $f(z)$ holomorphic on $B(0, R)$ with $R > 1$. Then, we can bound the singular values of a generic off-diagonal block $\tilde{C}$ in $f(A)$ with*

$$\sigma_l(\tilde{C}) \leqslant \gamma e^{-\frac{(\alpha + \alpha')l}{k}}, \quad \alpha = \log\left(\frac{R_A}{\rho}\right), \quad \alpha' = \log(R),$$

where $\gamma := \max_{|z|=R} |f(z)| \cdot \kappa_{max}^2 \cdot \|A\|_2 \cdot \Lambda(\rho, R_A, \mathcal{V}_A, R) \cdot \frac{k \cdot \rho}{R R_A - \rho}$ and κ_{max} is the maximum among the spectral condition numbers of the trailing submatrices of A.

Proof. Consider the partitioning $A = \left[\begin{smallmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{smallmatrix}\right]$ and for simplicity the case $k = 1$, $\bar{C} = uv^t$. The general case is obtained by linearity summing k objects of this kind coming from the SVD of $\bar{C}$ and applying Lemma 3.20. We rewrite the Dunford-Cauchy formula for $f(A)$

$$f(A) = \frac{1}{2\pi i} \int_{S^1} (zI - A)^{-1} f(z) dz.$$

Let $f(z) = \sum_{n \geqslant 0} a_n z^n$ be the Taylor expansion of $f(z)$ in $B(0, R)$. The corresponding off-diagonal block $\tilde{C}$ in $f(A)$ can be written as the outer product in Remark 2.2

$$\left[u \,\middle|\, \bar{D} \cdot u \,\middle|\, \ldots \,\middle|\, \bar{D}^{s-1} \cdot u\right] \cdot \left[\sum_{n=0}^{s-1} a_{n+1} (A^t)^n \bar{v} \,\middle|\, \ldots \,\middle|\, (a_s A^t + a_{s-1} I)\bar{v} \,\middle|\, a_s \bar{v}\right]^t [I \; 0]^t + g_s(A), \quad (8)$$

where $\bar{v} = [I \; 0]^t v$ and $g_s(A)$ is the remainder of the truncated Taylor series at order s. Since $f(z)$ is holomorphic in $B(0, R)$ the coefficients of $f(z)$ verify [23, Theorem 4.4c]

$$|a_j| \leqslant \max_{|z|=R} |f(z)| \cdot R^{-j}.$$

Applying Theorem 3.16 we get that $\forall r \in (\rho, R_A)$

$$\sigma_l(\tilde{C} - g_s(A)) \leqslant \gamma e^{-(\alpha+\alpha')l},$$

with $\alpha, \alpha', \delta, \kappa_{max}$ as in the thesis and $\gamma = \max_{|z|=R} |f(z)| \cdot \kappa_{max}^2 \|A\|_2 \cdot \Lambda(\rho, R_A, \mathcal{V}_A, R)$. Observing that this bound is independent on s and $\lim_{s\to\infty} g_s(A) = 0$ we get the thesis. □

Corollary 3.23. *Let $A \in \mathbb{C}^{m\times m}$ be a k-quasiseparable matrix, $z_0 \in \mathbb{C}$ and $R' \in \mathbb{R}^+$ such that $R'^{-1}(A - z_0 I)$ is enclosed in $(\rho, R_A, \mathcal{V}_A)$. Then, for any holomorphic function $f(z)$ in $B(z_0, R)$ with $R > R'$, any off-diagonal block $\tilde{C}$ in $f(A)$ has singular values bounded by*

$$\sigma_l(\tilde{C}) \leqslant \gamma e^{-\frac{(\alpha+\alpha')l}{k}}, \quad \alpha = \log\left(\frac{R_A}{\rho}\right), \quad \alpha' = \log\left(\frac{R}{R'}\right),$$

where $\gamma := \max_{|z-z_0|=R} |f(z)| \cdot \kappa_{max}^2 \cdot \|A - z_0 I\|_2 \cdot \Lambda(\rho, R_A, \mathcal{V}_A, R) \cdot \frac{k\cdot\rho}{RR_A - \rho R'}$ and κ_{max} is the maximum among the spectral condition numbers of the trailing submatrices of $R'^{-1}(A - z_0 I)$.

Proof. Define $g(z) = f(R'z + z_0)$ which is holomorphic on $B(0, \frac{R}{R'})$. Observing that $f(A) = g(R'^{-1}(A - z_0 I))$ we can conclude by applying Theorem 3.22. □

Remark 3.24. If we can find $z_0 \in \mathbb{C}$ such that $\|A - z_0 I\|_2 < R$ then it is always possible to find $(\rho, R_A, \mathcal{V}_A)$ with $\rho \in (0, 1)$ which satisfies the hypothesis of the previous corollary. A worst case estimate for $\frac{\rho}{R_A}$ is $\frac{\|A-z_0 I\|_2}{R}$ since this is the radius of a circle containing the spectrum of the rescaled matrix and — given that the Riemann map for a ball centered in 0 is the identity — $R_A = 1$.

Example 3.25 (Real spectrum). We here want to estimate the quantity $\frac{R_A}{\rho}$ in the case of a real spectrum for the matrix A. Suppose that — possibly after a scaling — the latter is contained in the symmetric interval $[-a, a]$ with $a \in (0, 1)$. The logarithmic capacity of this set is $\frac{a}{2}$ and the inverse of the associated Riemann map is $\psi(z) = z + \frac{a^2}{4}$. This follows by observing that the function $z + z^{-1}$ maps the circle of radius 1 into $[-2, 2]$, so then it is sufficient to compose the latter with two homothetic transformations to get $\psi(z)$. Moreover, observe that — given $r \geqslant \frac{a}{2}$ — ψ maps the circle of radius r into an ellipse of foci $[-a, a]$. Therefore, in order to get R_A it is sufficient to compute for which r we have $\psi(r) = 1$. This corresponds to finding the solution of $r + \frac{a^2}{4r} = 1$ which is greater than $\frac{a}{2}$. This yields

$$R_A = \frac{1 + \sqrt{1 - a^2}}{2} \quad \Rightarrow \quad \frac{R_A}{\rho} = \frac{1 + \sqrt{1 - a^2}}{a}.$$

4 Functions with singularities

If some singularities of f lie inside $B(z_0, R)$ then $f(A) \neq \int_{\partial B(z_0,R)} f(z)(zI - A)^{-1} dz$. However, since the coefficients of the Laurent expansion of f with negative degrees in (2) do not affect the result, the statement of Theorem 3.22 holds for the matrix $\int_{\partial B(z_0,R)} f(z)(zI - A)^{-1} dz$. In this section we prove that — under mild conditions — the difference of the above two terms still has a quasiseparable structure. This numerically preserves the quasiseparability of $f(A)$.

4.1 An extension of the Dunford-Cauchy integral formula

The main tool used to overcome difficulties in case of removable singularities will be the following result, which is an extension of the integral formula used in Definition 1.1.

Theorem 4.1. *Let $f(z)$ be a meromorphic function with a discrete set of poles $\mathcal{P}$ and $A \in \mathbb{C}^{m\times m}$ with spectrum $\mathcal{S}$ such that $\mathcal{S} \cap \mathcal{P} = \emptyset$. Moreover, consider Γ simple closed curve in the complex plane which encloses $\mathcal{S}$ and $T := \{z_1, \dots, z_t\} \subseteq \mathcal{P}$ subset of poles with orders $d_1, \dots, d_t$ respectively. Then*

$$\frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} f(z)dz = f(A) + \sum_{j=1}^{t} R_j(z_j I - A),$$

where R_j is the rational function

$$R_j(z) := \sum_{l=1}^{d_j} (-1)^{l+1} \frac{f_j^{(d_j-l)}(z_j)}{(d_j - l)!} z^{-l}$$

and $f_j(z) = (z - z_j)^{d_j} f(z)$, extended to the limit in z_j. In particular if the poles in T are simple then

$$\frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} f(z)dz = f(A) + \sum_{j=1}^{t} f_j(z_j) \cdot (z_j I - A)^{-1} = f(A) + \sum_{j=1}^{t} f_j(z_j)\mathfrak{R}(z_j).$$

Proof. We first prove the statement for A diagonalizable. Assume that $V^{-1}AV = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, then

$$\frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} f(z)dz = V^{-1} \begin{bmatrix} \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-\lambda_m} \end{bmatrix} V. \tag{9}$$

Applying the Residue theorem we arrive at

$$\frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z - \lambda_p} = \operatorname{Res}\left(\frac{f}{z-\lambda_p}, \lambda_p\right) + \sum_{j=1}^{t} \operatorname{Res}\left(\frac{f}{z-\lambda_p}, z_j\right), \qquad p = 1, \dots, m.$$

Since λ_p is a simple pole of $\frac{f}{z-\lambda_p}$ the first summand is equal to $f(\lambda_p)$.

On the other hand z_j is a pole of order d_j of $\frac{f}{z-\lambda_p}$, therefore its residue is

$$\operatorname{Res}\left(\frac{f}{z-\lambda_p}, z_j\right) = \frac{1}{(d_j-1)!}\lim_{z\to z_j}\frac{\partial^{d_j-1}}{\partial z^{d_j-1}}\left((z-z_j)^{d_j}\frac{f}{z-\lambda_p}\right) = \frac{1}{(d_j-1)!}\frac{\partial^{d_j-1}}{\partial z^{d_j-1}}\left(\frac{f_j}{z-\lambda_p}\right)(z_j).$$

One can prove by induction (see Appendix) that, given a sufficiently differentiable $f_j(z)$, it holds

$$\frac{\partial^{d-1}}{\partial z^{d-1}}\left(\frac{f_j(z)}{z-\lambda_p}\right) = \sum_{l=1}^{d}(-1)^{l+1}\frac{(d-1)!}{(d-l)!}f_j^{(d-l)}(z)(z-\lambda_p)^{-l}, \quad d \in \mathbb{N}. \tag{10}$$

Setting $d = d_j$ in (10) we derive

$$\operatorname{Res}\left(\frac{f}{z-\lambda_p}, z_j\right) = R_j(z_j - \lambda_p).$$

To conclude it is sufficient to rewrite the diagonal matrix in (9) as

$$\begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_m) \end{bmatrix} + \sum_{j=1}^{t} \begin{bmatrix} R_j(z_j-\lambda_1) & & \\ & \ddots & \\ & & R_j(z_j-\lambda_m) \end{bmatrix}.$$

We now prove the thesis for

$$A = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix},$$

because the general non diagonalizable case can be decomposed in sub-problems of that kind. We have that

$$\frac{1}{2\pi i}\int_\Gamma (zI-A)^{-1} f(z)dz = \frac{1}{2\pi i} \begin{bmatrix} \int_\Gamma \frac{f(z)}{z-\lambda} & \int_\Gamma \frac{f(z)}{(z-\lambda)^2} & \cdots & \int_\Gamma \frac{f(z)}{(z-\lambda)^m} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \int_\Gamma \frac{f(z)}{(z-\lambda)^2} \\ & & & \int_\Gamma \frac{f(z)}{z-\lambda} \end{bmatrix}.$$

In order to reapply the previous argument is sufficient to prove that

(i) $\mathrm{Res}(\frac{f}{(z-\lambda)^{h+1}}, \lambda) = \frac{f_j^{(h)}(\lambda)}{h!}$ $h = 1, \ldots, m-1$,

(ii) $\mathrm{Res}(\frac{f}{(z-\lambda)^{h+1}}, z_j) = \frac{R_j^{(h)}(z_j-\lambda)}{h!}$ $h = 1, \ldots, m-1$.

The point (i) is a direct consequence of the fact that λ is a pole of order $h+1$ of the function $\frac{f(z)}{(z-\lambda)^{h+1}}$. Concerning (ii) observe that z_j is again a pole of order d_j for the function $\frac{f(z)}{(z-\lambda)^{h+1}}$ so

$$\mathrm{Res}\left(\frac{f}{(z-\lambda)^{h+1}}, z_j\right) = \frac{1}{(d_j-1)!}\frac{\partial^{d_j-1}}{\partial z^{d_j-1}}\left(\frac{f_j(z)}{(z-\lambda)^{h+1}}\right)(z_j).$$

One can prove by induction (see Appendix) that, for each $d, h \in \mathbb{N}$:

$$\frac{\partial^{d-1}}{\partial z^{d-1}}\left(\frac{f_j(z)}{(z-\lambda)^{h+1}}\right) = \frac{(d-1)!}{h!}\sum_{l=1}^{d}(-1)^{l+h+1}\frac{(l+h-1)!}{(d-l)!(l-1)!} f_j^{(d-l)}(z)(z-\lambda)^{-(h+l)} \qquad (11)$$

Successive derivation of R_j repeated h times yields:

$$R_j^{(h)}(z) = \sum_{l=1}^{d_j}(-1)^{l+h+1}\frac{(l+h-1)!}{(d_j-l)!(l-1)!} f_j^{(d_j-l)}(z_j) z^{-(h+l)},$$

and by setting $d = d_j$ in (11) we finally get (ii). □

4.2 Functions with poles

As a direct application of Corollary 3.23 we can give a concise statement in the case of simple poles.

Corollary 4.2. *Let $A \in \mathbb{C}^{m\times m}$ be a quasiseparable matrix with rank k, $z_0 \in \mathbb{C}$ and $R' \in \mathbb{R}^+$ such that $R'^{-1}(A - z_0 I)$ is enclosed in $(\rho, R_A, \mathcal{V}_A)$. Consider $R > R'$ and a function $f(z)$ holomorphic on the annulus $\mathcal{A} := \{R' < |z - z_0| < R\}$. If the disc $B(z_0, R')$ contains t simple poles of f then any off-diagonal block $\tilde{C}$ in $f(A)$ has singular values bounded by*

$$\sigma_l(\tilde{C}) \leqslant \gamma e^{-\frac{(\alpha+\alpha')(l-tk)}{k}}, \quad \alpha = \log\left(\frac{R_A}{\rho}\right), \quad \alpha' = \log\left(\frac{R}{R'}\right),$$

where $\gamma := \max_{|z-z_0|=R} |f(z)| \cdot \kappa_{max}^2 \cdot \|A - z_0 I\|_2 \cdot \Lambda(\rho, R_A, \mathcal{V}_A, R) \cdot \frac{k\cdot\rho}{RR_A - \rho R'}$ and κ_{max} is the maximum among the spectral condition numbers of the trailing submatrices of $R'^{-1}(A - z_0 I)$.

Proof. Let $f(z) = \sum_{n\in\mathbb{Z}} a_n z^n$ be the series expansion of f in $\mathcal{A}$ and $z_1, \ldots, z_t$ be the simple poles of f inside $B(z_0, R')$. Then

$$|a_j| \leqslant \|f(z)\|_{\infty,\partial B(z_0,R)} \cdot \left(\frac{R'}{R}\right)^j, \qquad n \geqslant 0.$$

According to what we observed at the beginning of Section 4 we can apply Corollary 3.23 to the off-diagonal singular values of $B := \int_{\partial B(z_0,R')} f(z)(zI - A)^{-1}dz$. Moreover, using Theorem 4.1 we get

$$f(A) = B - \sum_{j=1}^{t} f_j(z_j) \cdot (z_j I - A)^{-1}.$$

Observing that the right summand has at most quasiseparable rank tk we can conclude, using Lemma 3.21, that the bound on the singular values of $f(A)$ is the same which holds for B, but shifted by the quantity $t \cdot k$. □

4.3 Functions with essential singularities

Consider the case of a function $f(z)$ holomorphic in $\mathbb{C} \setminus \{a\}$ with an essential singularity in a. Moreover, suppose that a is not an eigenvalue of the argument $A \in \mathbb{C}^{m\times m}$. In a suited punctured disk $B(a, R) \setminus \{a\}$ — which contains the spectrum of A — we can expand f as

$$f(z) := \sum_{n\in\mathbb{Z}} a_n (z - a)^n.$$

In particular we can decompose f as $f_1(z - a) + f_2((z - a)^{-1})$ with f_i holomorphic on $B(0, R)$ for $i = 1, 2$. Therefore

$$f(A) = f_1(A - aI) + f_2((A - aI)^{-1}).$$

Since f_1 and f_2 are both holomorphic and the operations of shift and inversion preserve the quasiseparable rank we can apply Theorem 3.22 and Lemma 3.20 in order to get estimates on the off-diagonal singular values of $f(A)$.

One can use this approach in the case of finite order poles and find equivalent bounds to Corollary 4.2, although in a less explicit form.

4.4 Functions with branches

We conclude this section describing how to re-adapt the approach in the case of functions with multiple branches. The same trick can be used to deal with other scenarios, such as the presence of singularities that has been described previously.

The main idea is that, in the integral definition of a matrix function, the path Γ does not need to be a single Jordan curve, but can be defined as a union of a finite number of them. The only requirement is that the function is analytic in the Jordan regions, and that the spectrum is contained in their union.

In our setting, it might happen that we cannot enclose the spectrum in a single ball without capturing also the branching point. However, it is always possible to cover it with the union of a finite number of such balls. In this context, assuming that the path Γ is split as the borders of t balls, denoted by $\Gamma_1, \dots, \Gamma_t$, one has

$$f(A) = \sum_{i=1}^{t} \int_{\Gamma_i} f(z)\mathfrak{R}(z)dz.$$

Assuming that the number t is small enough, we can obtain the numerical Quasiseparability of $f(A)$ by the quasiseparability of each of the addends and then relying on Lemma 3.20. Inside each $\Gamma_i = B(z_i, r_i)$ we can perform the change of variable $\tilde{z} := r_i(z - z_i)$ and write the resolvent as (here the coefficient D will be different by scaling and translation in every Γ_i):

$$\mathfrak{R}(\tilde{z}) = \begin{bmatrix} * & * \\ (\tilde{z}I - D)^{-1}C(\tilde{z})S_D(\tilde{z})^{-1} & * \end{bmatrix}, \qquad \begin{cases} (\tilde{z}I - D)^{-1} = \sum_{j\in\mathbb{Z}} D_j \tilde{z}^j \\ S_D^{-1}(\tilde{z}) = \sum_{s\in\mathbb{Z}} H_s \tilde{z}^s \end{cases}$$

The construction of the coefficients D_j can be done by writing D in Jordan canonical from as

$$V^{-1}DV = \begin{bmatrix} J_{\text{in}} & \\ & J_{\text{out}} \end{bmatrix}, \qquad V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}$$

where J_{in} refers to the part of the spectrum inside Γ_i, and J_{out} the one outside. Thanks to the change of variable in the integral, this corresponds to asking that the spectrum of J_{in} is inside the unit disc, and the one of J_{out} outside. Then, one has the following definition for D_j:

$$D_j = \begin{cases} V_1 J_{\text{in}}^{-j-1} W_1 & j < 0 \\ -V_2 J_{\text{out}}^{-j-1} W_2 & j \geqslant 0 \end{cases},$$

and an analogous formula holds for the coefficients H_s. This provides the Laurent expansion of the off-diagonal block in the integrand. A similar analysis to the one carried out in the previous sections can be used to retrieve the decay on the singular values of this block.

5 Computational aspects and validation of the bounds

In the previous sections we have proved that the numerical quasiseparable structure is often present in $f(A)$. This property can be used to speed up the matrix arithmetic operations and then to efficiently evaluate $f(A)$ by means of contour integration. We briefly describe the strategy in the next subsections and we refer the reader to [22] for more details. In Section 5.3 we will compare our bounds with the actual decay in some concrete cases.

5.1 Representation and arithmetic operations

In order to take advantage of the quasiseparable structure we need a representation that enable us to perform the storage and the matrix operations cheaply. We rely on the framework of Hierarchical representations originally introduced by Hackbusch [21,22] in the context of integral and partial differential equations. It consists in a class of recursive block representations with structured sub-matrices that allows the treatment of a number of data-sparse patterns. Here, we consider a particular member of this family — sometimes called Hierarchical off-diagonal low-rank representation (HODLR) — which has a simple formulation and an effective impact in handling quasiseparable matrices.

Let $A \in \mathbb{C}^{m\times m}$ be a k-quasiseparable matrix and consider the partitioning

$$A = \left[\begin{smallmatrix} A_{11} & A_{22} \\ A_{21} & A_{22} \end{smallmatrix}\right],$$

where $A_{11} \in \mathbb{C}^{m_1\times m_1}$, $A_{22} \in \mathbb{C}^{m_2\times m_2}$, with $m_1 := \lfloor \frac{m}{2} \rfloor$ and $m_2 := \lceil \frac{m}{2} \rceil$. Observe that the antidiagonal blocks A_{12} and A_{21} do not involve any element of the main diagonal of A, hence we can represent them in a compressed form as an outer product of rank k. Moreover, the diagonal blocks A_{11} and A_{22} are square matrices which are again k-quasiseparable. Therefore it is possible to re-apply this procedure recursively. We stop when the diagonal blocks reach a minimal dimension $m_{\min}$, and we store them as full matrices. The process is described graphically in Figure 1.

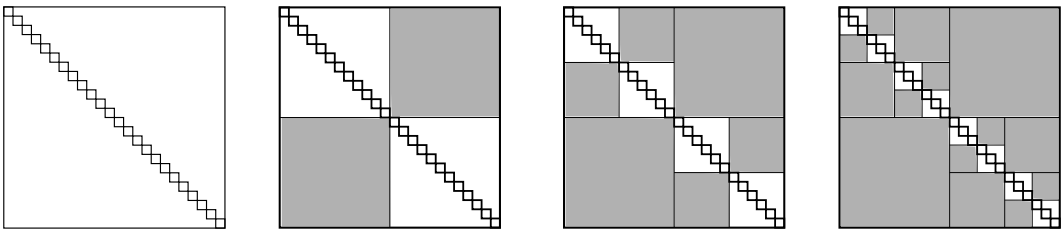

Figure 1: The behavior of the block partitioning in the HODLR-matrix representation. The blocks filled with grey are low rank matrices represented in a compressed form, and the diagonal blocks in the last step are stored as dense matrices.

If $m_{\min}$ and k are negligible with respect to m then the storage cost of each sub-matrix is $O(m)$. Since the levels of the recursion are $O(\log(m))$, this yields a linear-polylogarithmic memory consumption with respect to the size of the matrix.

The HODLR representation acts on a matrix by compressing many of its sub-blocks. Therefore, it is natural to perform the arithmetic operations in a block-recursive fashion. The basic steps of these procedures require arithmetic operations between low-rank matrices or $m_{min} \times m_{min}$-matrices. If the rank of the off-diagonal blocks is small compared to m, then the algorithms performing the arithmetic operations have linear polylogarithmic complexities [9][Chapter 6]. The latter are summarized in Table 1 where it is assumed that the constant k bounds the quasiseparable rank of all the matrices involved. Moreover, the operations are performed adaptively with respect to the rank of the blocks. This means that the result of an arithmetic operation will be an HODLR matrix with the same partitioning, where each low rank block is a truncated reduced SVD of the corresponding block of the exact result. This operation can be carried out with linear cost, assuming the quasiseparable stays negligible with respect to m. Hence the rank is not fixed a priori but depends on a threshold ϵ at which the truncation is done. We refer to [22] for a complete description. In our experiments we set ϵ equal to the machine precision $2.22 \cdot 10^{-16}$ and $m_{\min} = 64$.

Operation	Computational complexity
Matrix-vector multiplication	$O(km\log(m))$
Matrix-matrix addition	$O(k^2m\log(m))$
Matrix-matrix multiplication	$O(k^2m\log(m)^2)$
Matrix-inversion	$O(k^2m\log(m)^2)$
Solve linear system	$O(k^2m\log(m)^2)$

Table 1: Computational complexity of the HODLR-matrix arithmetic. The operation *Solve linear system* comprises to compute the LU factorization of the coefficient matrix and to solve the two triangular linear systems

5.2 Contour integration

The Cauchy integral formula (1) can be used to approximate $f(A)$ by means of a numerical integration scheme. Recall that, given a complex valued function $g(x)$ defined on an interval $[a,b]$ one can approximate its integral by

$$\int_a^b g(x)dx \;\approx\; \sum_{k=1}^{N} w_k \cdot g(x_k) \tag{12}$$

where w_k are the *weights* and x_k are the *nodes*. Since we are interested in integrating a function on S^1 we can write

$$\frac{1}{2\pi i}\int_{S^1} f(z)(zI-A)^{-1}dz = \frac{1}{2\pi}\int_0^{2\pi} f(e^{ix})(I-e^{-ix}A)^{-1}dx,$$

where we have parametrized S^1 by means of e^{ix}. The right-hand side can be approximated by means of (12), so we obtain:

$$f(A) \approx \frac{1}{2\pi}\sum_{k=1}^{N} w_k \cdot f(e^{ix_k})(I-e^{-ix_k}A)^{-1} = \frac{1}{2\pi}\sum_{k=1}^{N} w_k \cdot e^{ix_k} f(e^{ix_k})\mathfrak{R}(e^{ix_k}). \tag{13}$$

This approach has already been explored in [18], mainly for the computation of $f(A)b$ due to the otherwise high cost of the inversions in the general case. The pseudocode of the procedure is reported in Algorithm 1.

Algorithm 1 — based on (13) — can be carried out cheaply when A is represented as an HODLR-matrix, since the inversion only requires $O(m\log^2(m))$ flops. Moreover, not only the resolvent $\mathfrak{R}(e^{ix_k})$ is representable as a HODLR-matrix, but the same holds for the final result $f(A)$ in view of Theorem 3.22. This guarantees the applicability of the above strategy even when dealing with large dimensions.

The results in Section 4 enable us to deal with functions having poles inside the domain of integration. The only additional step that is required is to compute the correction term described in Theorem 4.1. Notice that this step just requires additional evaluations of the resolvent and so does not change the asymptotic complexity of the whole procedure.

We show now an example where Theorem 4.1 can be used to derive an alternative algorithm for the evaluation of matrix functions with poles inside the domain.

Algorithm 1 Pseudocode for the evaluation of a contour integral on S^1

procedure ContourIntegral(f, A) ▷ Evaluate $\frac{1}{2\pi i}\int_{S^1} f(z)(zI-A)^{-1}dz$
  $N \leftarrow 1$
  $M \leftarrow f(1)\cdot(I-A)^{-1}$
  $err \leftarrow \infty$
  while $err > \sqrt{u}$ **do**
    $M_{\text{old}} \leftarrow M$
    $M \leftarrow \frac{1}{2}M$ ▷ The new weights are applied to the old evaluations
    $N \leftarrow 2N$
    for $j = 1, 3, \ldots, N-1$ **do** ▷ Sum the evaluations on the new nodes
      $z \leftarrow e^{\frac{2\pi i j}{N}}$
      $M \leftarrow M + \frac{zf(z)}{N}\cdot(zI-A)^{-1}$
    end for
    $err \leftarrow \|M - M_{\text{old}}\|_2$
  end while
  return M
end procedure

More precisely, we consider a matrix A with spectrum contained in the unit disc, and the evaluation of the matrix function $f(A)$ with $f(z) = \frac{e^z}{\sin(z)}$. Applying Theorem 4.1 yields

$$f(A) = \int_{S^1} f(z)\mathfrak{R}(z)dz + A^{-1}.$$

One can then choose to obtain $f(A)$ by computing $e^A \cdot (\sin A)^{-1}$, which requires the evaluation of two integrals and one inverse, or using the above formula, which only requires one integral and an inverse.

We used an adaptive doubling strategy for the number of nodes i.e., starting with N-th roots of the unit for a small value of N. We apply the quadrature rule (13) and we double N until the quality of the approximation is satisfying. In order to check this, we require that the norm of the difference between two consecutive approximations is smaller than a certain threshold. The 2-norm of an HODLR-matrix can be estimated in linear time as shown in [22]. Since the quadrature rule is quadratically convergent [31] and the magnitude of the distance between the approximations at step k and $k+1$ is a heuristic estimate for the error at step k we choose as threshold $\sqrt{u}$ where u is the unit round-off. In this way we should get an error of the order of u.

We show in Table 2, where the approach relying on Theorem 4.1 and on computing the function separately are identified by the labels "sum" and "inv", respectively, that the first choice is faster (due to the reduced number of inversions required) and has a similar accuracy. The matrices in this example have been chosen to be 1-quasiseparable and Hermitian, and we have verified the accuracy of the results by means of a direct application of Definition 1.1. In particular, the timings confirm the almost linear complexity of the procedure.

5.3 Validation of the bounds

This section is devoted to check the accuracy of the estimates for the singular values that we have proved in the paper. In order to do so we compute some matrix function on quasiseparable matrices and verify the singular values decay in one large off-diagonal block. In particular, for a matrix of order m — m even — we consider the off-diagonal block with row indices from

Size	t_{inv}	Res_{inv}	t_{sum}	Res_{sum}
128	2.95 s	$1.33 \cdot 10^{-13}$	1.51 s	$3.3 \cdot 10^{-14}$
256	9.78 s	$4.58 \cdot 10^{-12}$	4.84 s	$1.2 \cdot 10^{-12}$
512	24.6 s	$5.55 \cdot 10^{-11}$	12.2 s	$3.02 \cdot 10^{-12}$
1,024	57 s	$5.87 \cdot 10^{-11}$	23.5 s	$3.92 \cdot 10^{-11}$
2,048	132 s	$6.01 \cdot 10^{-11}$	48.1 s	$3.99 \cdot 10^{-11}$
4,096	245 s	$6.59 \cdot 10^{-11}$	127 s	$5.69 \cdot 10^{-10}$

Table 2: Timing and accuracy on the computation of the matrix function $f(z) = e^z \sin(z)^{-1}$ on a 1-quasiseparable Hermitian matrix A with spectrum contained the unit disc. The residues are measured relatively to the norm of the computed matrix function $f(A)$.

$\frac{m}{2} + 1$ to m and column indices from 1 to $\frac{m}{2}$. Then, we compare the obtained result with the theoretical bound coming from Theorem 3.22. Notice that Theorem 3.22 provides a family of bounds depending on a parameter R which can be chosen as long as $f(z)$ is holomorphic in $B(0, R)$. So, in every experiment we estimated the l-th singular value by choosing the parameter R which provides the tighter bound, among the admissible values for the function f under consideration.

We choose two particular classes of 1-quasiseparable matrices for the tests, since we can easily determine the bounds on them:

Hermitian tridiagonal matrices These matrices are generated with elements taken from a random Gaussian distribution $N(0, 1)$, and are then scaled and shifted so that their spectrum is contained in a ball of center 0 and radius $\frac{3}{4}$. These matrices are normal and the same holds for their submatrices, so we can avoid the computation of the constants $\kappa_s(\cdot)$ which are all equal to 1.

Hessenberg (scaled) unitary matrices We consider a random unitary matrix which is also upper Hessenberg, and so in particular it is 1-quasiseparable (since unitary matrices are rank symmetric - the rank of the lower off-diagonal blocks is equal to the corresponding block above). We then scale the matrices multiplying by $\frac{3}{4}$, in order to keep the spectrum on the circle of radius $\frac{3}{4}$. We obtain these matrices in MATLAB by running the command `[A,~] = .75 * qr(hess(randn(N)));` where N is the chosen dimension.

As a first example we consider the matrix exponential e^A which can be easily computed by means of `expm`. We have computed it for many random tridiagonal matrices of size 1000×1000, and the measured and theoretical decays in the submatrix $e^A(501 : 1000, 1 : 500)$ are reported in Figure 2.

Similarly, in Figure 3 we have reported the analogous experiment concerning the function $\log(4I + A)$. In fact, in order for the logarithm to be well defined, we need to make sure that the spectrum of the matrix inside the logarithm does not have any negative value.

As a last example for the tridiagonal matrices we have considered the case of the function $\sqrt{4I + A}$, where the matrix has been shifted again in order to obtain a reasonable estimate by moving the spectrum away from the branching point. The result for this experiment is reported in Figure 4.

In the same figures we have reported also the experiments in the case of the scaled unitary Hessenberg matrix. In this case the variance in the behavior of the singular values was very small in the experiments, and so we have only reported one example for each case.

Notice that while in the symmetric (or Hermitian) case every trailing diagonal submatrix is guaranteed to be normal, this is not true anymore for the scaled unitary Hessenberg matrices.

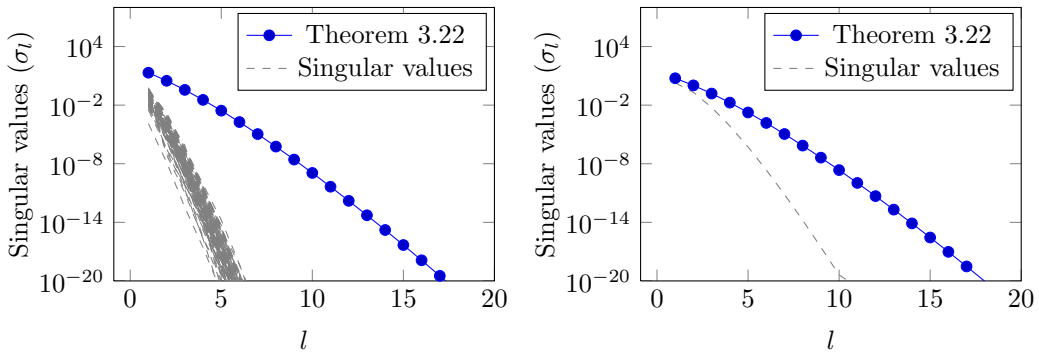

Figure 2: On the left, the bound on the singular values of the off-diagonal matrices of e^A for 100 random Hermitian tridiagonal matrices scaled in order to have spectral radius $\frac{3}{4}$ are shown. In the right picture the same experiment with a scaled upper Hessenberg unitary matrix is reported (with 1 matrix only).

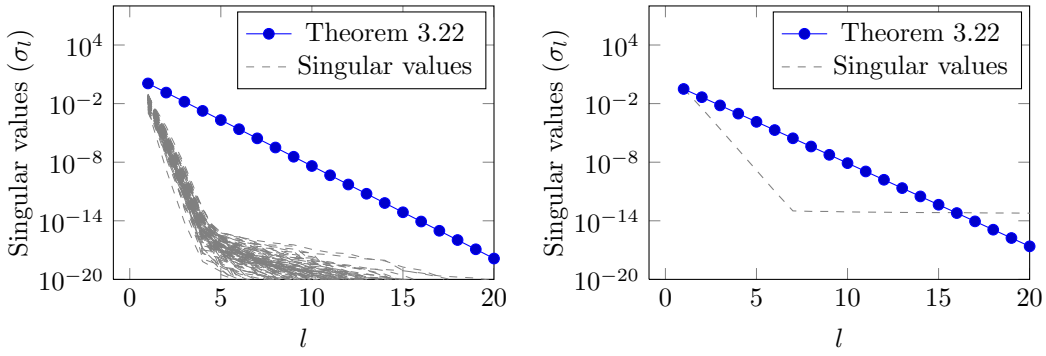

Figure 3: The picture reports the same experiment of Figure 2, with the logarithm in place of the exponential. The matrices have however been shifted by $4I$ in order to make the function well-defined. Since this corresponds to evaluating the function $\log(z+4)$ on the original matrix, one can also find a suitable ball centered in 0 where the function is analytic.

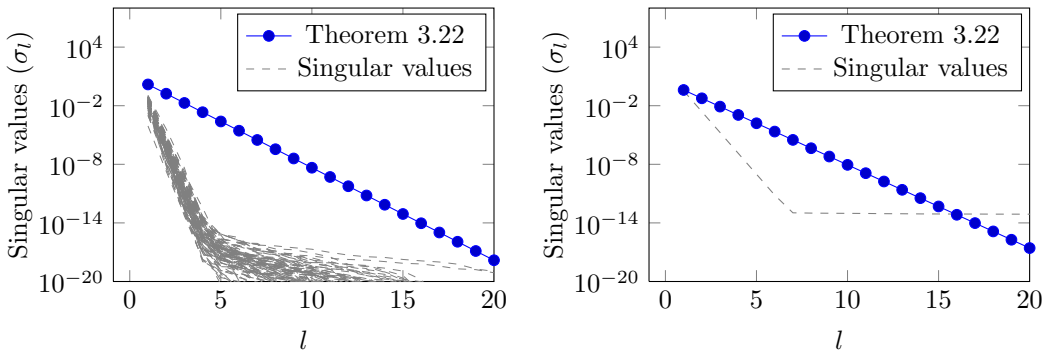

Figure 4: In the left picture the bounds on the singular values of the off-diagonal matrices of $\sqrt{4I + A}$ for 100 random Hermitian tridiagonal matrix scaled in order to have spectral radius $\frac{3}{4}$ are shown. In the right picture the same experiment is repeated for a scaled and shifted upper Hessenberg unitary matrix.

Nevertheless, one can verify in practice that these matrices are still not far from normality, and so the bounds that we obtain do not degrade much.

6 Concluding remarks

The numerical preservation of the quasiseparable structure when computing a matrix function is an evident phenomenon. Theoretically, this can be explained by the existence of accurate rational approximants of the function over the spectrum of the argument. In this work we have given a closer look to the off-diagonal structure of $f(A)$ providing concrete bounds for its off-diagonal singular values. The off-diagonal blocks have been described as a product between structured matrices with a strong connection with Krylov spaces. This —combined with polynomial interpolation techniques— is the key for proving the bounds.

Moreover, we have developed new tools to deal with the difficulties arising in the treatment of singularities and branching points. In particular the formula of Corollary 4.2 can be employed with the technology of Hierarchical matrices for efficiently computing matrix functions with singularities. An example of this strategy has been provided along with the numerical validation of the bounds.

A Appendix

Proposition A.1. *Let $f \in \mathcal{C}^{\infty}(\mathbb{C})$ and $\lambda \in \mathbb{C}$ then*

$$\frac{\partial^{d-1}}{\partial z^{d-1}}\left(\frac{f(z)}{(z-\lambda)^{h+1}}\right) = \frac{(d-1)!}{h!}\sum_{l=1}^{d}(-1)^{l+h+1}\frac{(l+h-1)!}{(d-l)!(l-1)!}f^{(d-l)}(z)(z-\lambda)^{-(h+l)}, \qquad \forall d \in \mathbb{N}^{+}, h \in \mathbb{N}.$$

Proof. For every fixed $h \in \mathbb{N}$ we proceed by induction on d. For $d = 1$ we get

$$\frac{f(z)}{(z-\lambda)^{h+1}} = \frac{0!}{h!}(-1)^{2}\frac{h!}{0!0!}\frac{f(z)}{(z-\lambda)^{h+1}}.$$

For the inductive step, let $d > 1$ and observe that

$$\begin{aligned}
\frac{\partial^d}{\partial z^d}\left(\frac{f(z)}{(z-\lambda)^{h+1}}\right) &= \frac{\partial}{\partial z}\left(\frac{\partial^{d-1}}{\partial z^{d-1}}\left(\frac{f(z)}{(z-\lambda)^{h+1}}\right)\right) \\
&= \frac{\partial}{\partial z}\left(\frac{(d-1)!}{h!}\sum_{l=1}^{d}(-1)^{l+h+1}\frac{(l+h-1)!}{(d-l)!(l-1)!}f^{(d-l)}(z)(z-\lambda)^{-(h+l)}\right) \\
&= \frac{(d-1)!}{h!}\sum_{l=1}^{d}(-1)^{l+h+1}\frac{(l+h-1)!}{(d-l)!(l-1)!}f^{(d+1-l)}(z)(z-\lambda)^{-(h+l)} \\
&+ \frac{(d-1)!}{h!}\sum_{l=1}^{d}(-1)^{l+h+2}(h+l)\frac{(l+h-1)!}{(d-l)!(l-1)!}f^{(d-l)}(z)(z-\lambda)^{-(h+l+1)} \\
&= \frac{(d-1)!}{h!}\sum_{l=1}^{d}(-1)^{l+h+1}\frac{(l+h-1)!}{(d-l)!(l-1)!}f^{(d+1-l)}(z)(z-\lambda)^{-(h+l)} \\
&+ \frac{(d-1)!}{h!}\sum_{l=2}^{d+1}(-1)^{l+h+1}(h+l-1)\frac{(l+h-2)!}{(d+1-l)!(l-2)!}f^{(d+1-l)}(z)(z-\lambda)^{-(h+l)} \\
&= \frac{d!}{h!}\sum_{l=1}^{d+1}(-1)^{l+h+1}\frac{(l+h-1)!}{(d+1-l)!(l-1)!}f^{(d+1-l)}(z)(z-\lambda)^{-(h+l)}.
\end{aligned}$$

□

References

[1] M. Benzi and P. Boito. Decay properties for functions of matrices over C^*-algebras. *Linear Algebra Appl.*, 456:174–198, 2014.

[2] M. Benzi, P. Boito, and N. Razouk. Decay properties of spectral projectors with applications to electronic structure. *SIAM Rev.*, 55(1):3–64, 2013.

[3] M. Benzi and V. Simoncini. Decay bounds for functions of Hermitian matrices with banded or Kronecker structure. *SIAM J. Matrix Anal. Appl.*, 36(3):1263–1282, 2015.

[4] D. A. Bini, B. Iannazzo, and B. Meini. *Numerical Solution of Algebraic Riccati Equations*. Fundamentals of Algorithms n. 9. SIAM, Philadelphia, 2012.

[5] D. A. Bini, G. Latouche, and B. Meini. *Numerical methods for structured Markov chains*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2005. Oxford Science Publications.

[6] D. A. Bini, S. Massei, and L. Robol. Efficient cyclic reduction for Quasi-Birth–Death problems with rank structured blocks. *Appl. Numer. Math.*, 2016.

[7] D. A. Bini, S. Massei, and L. Robol. On the decay of the off-diagonal singular values in cyclic reduction. *arXiv preprint arXiv:1608.01567*, 2016.

[8] D. A. Bini and B. Meini. The cyclic reduction algorithm: from Poisson equation to stochastic processes and beyond. In memoriam of Gene H. Golub. *Numer. Algorithms*, 51(1):23–60, 2009.

[9] S. Börm, L. Grasedyck, and W. Hackbusch. Hierarchical matrices. *Lecture notes*, 21:2003, 2003.

[10] B. L. Buzbee, G. H. Golub, and C. W. Nielson. On direct methods for solving Poisson's equations. *SIAM J. Numer. Anal.*, 7:627–656, 1970.

[11] C. Canuto, V. Simoncini, and M. Verani. On the decay of the inverse of matrices that are sum of Kronecker products. *Linear Algebra Appl.*, 452:21–39, 2014.

[12] S. Chandrasekaran, P. Dewilde, M. Gu, and N. Somasunderam. On the numerical rank of the off-diagonal blocks of Schur complements of discretized elliptic PDEs. *SIAM J. Matrix Anal. Appl.*, 31(5):2261–2290, 2010.

[13] M. Crouzeix. Numerical range and functional calculus in Hilbert space. *J. Funct. Anal.*, 244(2):668–690, 2007.

[14] Y. Eidelman and I. Gohberg. On generators of quasiseparable finite block matrices. *Calcolo*, 42(3-4):187–214, 2005.

[15] Y. Eidelman, I. Gohberg, and I. Haimovici. *Separable type representations of matrices and fast algorithms. Vol. 1*, volume 234 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer, Basel, 2014. Basics. Completion problems. Multiplication and inversion algorithms.

[16] S. W. Ellacott. Computation of Faber series with application to numerical polynomial approximation in the complex plane. *Math. Comp.*, 40(162):575–587, 1983.

[17] F. R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998. Translated from the Russian by K. A. Hirsch, Reprint of the 1959 translation.

[18] I. P. Gavrilyuk, W. Hackbusch, and B. N. Khoromskij. $\mathscr{H}$-matrix approximation for the operator exponential with applications. *Numer. Math.*, 92(1):83–111, 2002.

[19] I. P. Gavrilyuk, W. Hackbusch, and B. N. Khoromskij. Data-sparse approximation to a class of operator-valued functions. *Math. Comp.*, 74(250):681–708, 2005.

[20] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[21] W. Hackbusch. A sparse matrix arithmetic based on $\mathscr{H}$-matrices. Part I: Introduction to $\mathscr{H}$-matrices. *Computing*, 62(2):89–108, 1999.

[22] W. Hackbusch. *Hierarchical matrices: algorithms and analysis*, volume 49 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2015.

[23] P. Henrici. *Applied and computational complex analysis. Vol. 1.* Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988.

[24] N. J. Higham. *Functions of matrices*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. Theory and computation.

[25] R. W. Hockney. A fast direct solution of Poisson's equation using Fourier analysis. *J. Assoc. Comput. Mach.*, 12:95–113, 1965.

[26] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[27] P. Lancaster and M. Tismenetsky. *The theory of matrices*. Computer Science and Applied Mathematics. Academic Press, Inc., Orlando, FL, second edition, 1985.

[28] N. S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, New York-Heidelberg, 1972. Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180.

[29] A. I. Markushevich. *Theory of functions of a complex variable. Vol. I, II, III*. Chelsea Publishing Co., New York, english edition, 1977. Translated and edited by Richard A. Silverman.

[30] Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.

[31] L. N. Trefethen and J. Weideman. The exponentially convergent trapezoidal rule. *SIAM Rev.*, 56(3):385–458, 2014.

[32] R. Vandebril, M. Van Barel, and N. Mastronardi. *Matrix computations and semiseparable matrices. Linear systems*, volume 1. Johns Hopkins University Press, Baltimore, MD, 2008.

[33] R. Vandebril, M. Van Barel, and N. Mastronardi. *Matrix computations and semiseparable matrices. Eigenvalue and singular value methods*, volume 2. Johns Hopkins University Press, Baltimore, MD, 2008.